# Abundance Profile Search

Purpose. Selection of protein families (COGs and Pfams) in isolate genomes, metagenomes or metagenome bins (or workspace scaffold sets) based on their relative abundance; similar to Phylogenetic Profiler for gene selection, but operates on protein families rather than individual genes.

Menu Navigation. Compare Genomes → Abundance Profiles → Abundance Profile Search
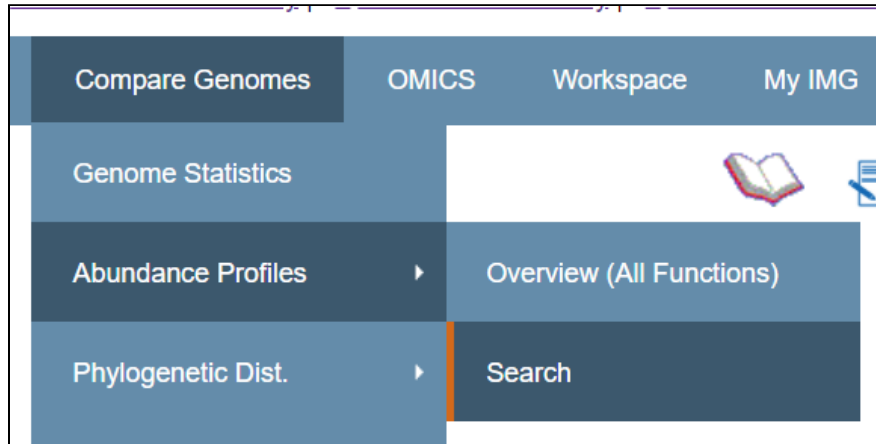
Functionality. Select the type of protein families (COG or Pfam), normalization method, and display of results. Abundance cut-offs can be set up for the genomes/metagenomes/bins of interest (e.g., – find all COGs in Ferroplasma Type I that are at least twice as abundant as in Ferroplasma Type II and are at least twice less abundant than in *Thermoplasmatales archaeon*). The families in the results table can be selected and added to the Function Cart, while gene counts in the table are linked to the corresponding lists of genes, which can be also selected and added to the Gene Cart.

Abundance Profile Search Form

Data Set selection

Selecting Genomes:
Click on one or more genome names, and then click the "Add" button to confirm selection.

NEW Scaffold Set Selector:
Click the **"Upload Sets"** to view a select scaffold sets dialog.

My Workspace - Scaffold Sets

Please select some scaffold sets:

Select	File Name	Num. of Scaffolds	Comment	My Lineage	File Size
☐	Butyrivibrio	4	Butyrivibrio proteoclasticus B316 abudnace test		387 B
☐	from_gene_json	1			311 B
☑	Freshwater_sediment_methanotrophic	364	bin 1 Freshwater sediment methanotrophic		17.368 KB
☐	MT_N_pt4	13716			629.822 KB
☐	json_union	4			398 B
☐	json_test_02	4			398 B

Cancel   OK

Less Abundant Than In   1
Freshwater sediment methanotrophic microbial comn

Real time vs Background Computation

Pressing the "Go" button runs the computation in real time with some limitations:
1. Selecting a maximum of 100 data sets.
2. Scaffold set size should have less than 20 scaffolds.

Selecting too many data sets may result in timeouts.

We recommend **"Background Computation"** which is only available in IMG/MER.
Using background computation you can select more data sets and larger scaffold sets. Your results will be saved in your workspace.

Functional Classification:
- ● COG
- ○ Enzyme
- ○ KO
- ○ Pfam
- ○ Tigrfam

Normalization Method:
- ● None
- ○ Frequency

Show Results As:
- ☑ Count
  - ● Gene count
  - ○ Estimated gene copies

More Abundant Cut-Off [1 ▾]   Less Abundant Cut-Off [1 ▾]

Enter matching text for highlighting clusters/rows. (E.g., "kinase".)

[   ]

[Go]   [Reset]

Real time computation

Submit as Computation Job

Submit a computation to run in the background.

- ● Save as a new job with name: [abundance_test1]
- ○ Replace the selected job: [▾]

Comment: [   ]

[Submit Computation]   **Submit for background computation (recommended)**

Output Results

Real time and background job output results will be identical.

Abundance Profile Search Results

1 - Butyrivibrio fibrisolvens 16/4
2 - Butyrivibrio fibrisolvens DSM 3071
3 - Butyrivibrio
4 - Abundance_test_merfs
5 - Freshwater sediment methanotrophic microbial communities from Lake Washington under simulated oxygen tension - Sediment Metagenome 20_HOW5 (SPAdes)

[pink box] Less abundant
[green box] More abundant

Normalization Method: **None**

| Add Selected to Function Cart | Select All | Clear All |

Filter column: COG ID ▾ Filter text ▾ : [   ] Apply ⓘ

Export Page 1 of 36 << first < prev **1** 2 3 4 5 6 7 8 9 10 next > last >> 100 ▾

| Column Selector | Select Page | Deselect Page |

Selection	COG ID ▲	COG Name	Butyrivibr fibrisolve 16 650377919	Butyrivibr fibrisolve DSM 2585428068	Butyrivibrio	Abundance_test_merfs	Freshwater sediment methanotro 3300024970
☐	COG0001	Glutamate-1-semialdehyde aminotransferase	0	0	0	0	59
☐	COG0002	N-acetyl-gamma-glutamylphosphate reductase	1	0	0	0	58
☐	COG0003	Anion-transporting ATPase, ArsA/GET3 family	0	0	0	0	21

Notes. Abundance Profile Search does not take into account the degree of assembly of a metagenome, i.e., the differences in read depth coverage between the genes and gene families in well assembled and poorly assembled metagenomes. Beware when comparing poorly assembled metagenomes with well assembled metagenomes.

Workspace Job Output example

For background computation you will receive an email when your results have been completed. All results will be saved in your workspace.

Workspace Menu navigation:
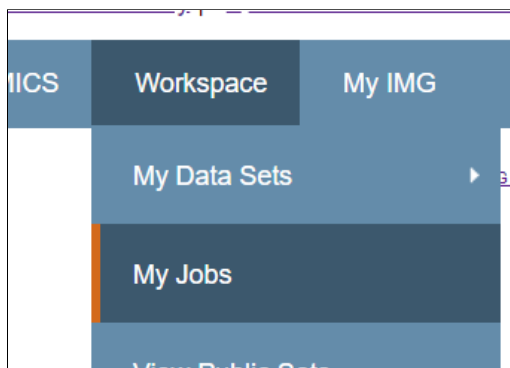
Job summary list:

List of all my jobs.

Program ⇕	Description ⇕	Count ⇕
Abundance Profile Overview		10
Abundance Profile Search		2

Abundance Profile Search list:

Clicking the **"2"** from above, you will get all of your "Abundance Profile Search" jobs.
Click the **"completed"** url to view your results.

Select	Name ▲	Type	Start Time	Parameters	Comment	File Size	End Time	Status
☐	abundance_test1	Abundance Profile Search	2021/03/31 10:45:21		abundance test1 norma value unchecked	446.645 KB	2021/04/16 21:12:34	completed

Results page:

Abundance Profile Search Results

1 - Butyrivibrio fibrisolvens 16/4
2 - Butyrivibrio fibrisolvens DSM 3071
3 - Butyrivibrio
4 - Abundance_test_merfs
5 - Freshwater sediment methanotrophic microbial communities from Lake Washington under simulated oxygen tension - Sediment Metagenome 20_HOW5 (SPAdes)

▨	Less abundant
▨	More abundant

Normalization Method: **None**

[Add Selected to Function Cart] [Select All] [Clear All]

Filter column: [COG ID ▾] Filter [text ▾] : [_____] [Apply] ❓

[Export] Page 1 of 36 << first < prev **1** 2 3 4 5 6 7 8 9 10 next > last >> [100 ▾]

[Column Selector] [Select Page] [Deselect Page]

Selection	COG ID ▲	COG Name	Butyrivibr fibrisolve 16 650377919	Butyrivibr fibrisolve DSM 2585428068	Butyrivibrio	Abundance_test_merfs	Freshwater sediment methanotro 3300024970
☐	COG0001	Glutamate-1-semialdehyde aminotransferase	0	0	0	0	59
☐	COG0002	N-acetyl-gamma-glutamylphosphate reductase	1	0	0	0	58
		Anion-transporting ATPase, ArsA/GET3					